

WHAT IS CLAIMED IS:

1. A method of voice recognition in a noise-ridden acoustic signal comprising:

a phase of digitizing and subdividing the noise-ridden acoustic signal into a sequence of temporal frames,

a phase of parametrization of speech-containing temporal frames so as to obtain a vector of parameters, per frame, in the frequency domain, this vector of parameters expressing the acoustic contents of the frame,

a shape-recognition phase in which the vectors of parameters are assessed with respect to references pre-recorded in a reference space during a preliminary learning phase, so as to obtain recognition by the determining of at least one reference which is closest to the vector of parameters,

a phase of reiterative searching for successive noise models in the sequence of temporal frames, a new noise model replacing a current noise model, a noise model comprising several successive frames,

wherein the method comprises:

a phase of searching for a noise transition between the new noise model and the current model,

and wherein, when the noise transition has been detected, the method comprises a phase of updating the reference space as a function of the new noise model, the parametrization phase including a step of matching the parameters to the new noise model.

2. A method of voice recognition according to claim 1, wherein the phase of searching for a noise transition comprises a step of searching for an energy incompatibility and/or a step of searching for a spectral incompatibility between the new noise model and the current model, the detection of an incompatibility expressing a noise transition.

3. A method of voice recognition according to claim 2, wherein the step of searching for an energy incompatibility comprises the comparison of the ratio between the mean energy E_{newmod} of the new noise model and the mean energy of the current noise model E_{modcurr} with a low threshold S' and a high threshold S , an energy incompatibility being found if the ratio is outside the interval delimited by the two thresholds S, S' .

00113-050901

4. A method of voice recognition according to claim 3, wherein the step of searching for an energy incompatibility also comprises a comparison of the mean energy E_{newmod} of the new noise model and the mean energy of the current noise model E_{modcurr} with an energy floor threshold E_{min} below which the noise is negligible, the energy incompatibility determined by the comparison of the ratio between the mean energy of the new noise model E_{newmod} and the mean energy of the current noise model E_{modcurr} being ignored when the mean energy of the new noise model E_{newmod} and the mean energy of the current noise model E_{modcurr} are both below the energy floor threshold E_{min} .

5. A method of voice recognition according to claim 2, wherein the step of searching for spectral incompatibility comprises, on the basis of spectral coefficients $B_{i,\text{modcurr}}$ $B_{i,\text{newmod}}$ respectively expressing the spectral energy of the frames of the current noise model and the spectral energy of the frames of the new noise model in at least one frequency channel i , a comparison of the ratio between the spectral coefficient $B_{i,\text{newmod}}$ associated with the frequency channel i of the new noise model and the spectral coefficient $B_{i,\text{modcurr}}$ associated with the same frequency channel i of the current noise model with a low threshold Sf' and a high threshold Sf , a spectral incompatibility being found if the ratio is located outside the interval delimited by the two thresholds, Sf , Sf' .

6. A method of voice recognition according to claim 5, wherein the step of searching for a spectral incompatibility also comprises, for at least one frequency channel i , a comparison of the spectral coefficient $B_{i,\text{newmod}}$ of the new noise model in this frequency channel i and of the spectral coefficient $B_{i,\text{modcurr}}$ of the current noise model in this frequency channel i with a floor spectral coefficient $B_{i,\text{min}}$ associated with this frequency channel i , namely a floor below which the noise is negligible, a spectral incompatibility determined by the comparison of the ratio between spectral coefficients being ignored when, for this frequency channel i , the spectral coefficients of the new model and of the current model are both below the floor spectral coefficient $B_{i,\text{min}}$.

7. A method of voice recognition according to claim 1, wherein the parametrization phase comprises a step of determining spectral coefficients $B_{i,\text{par}}$ each associated with a frequency channel i each expressing a

09851133 050901

A

representation of the spectral energy of a frame containing speech in the frequency channel i ,

the parameter-matching step comprising a determining, for each spectral coefficient $B_{i,par}$, of a robustness operator $OpRob(B_{i,par})$, this robustness

5 operator expressing the confidence to be attached to the spectral coefficient $B_{i,par}$ with respect to the noise level of the new noise model in the same frequency channel i ,

a weighting of the spectral coefficient $B_{i,par}$ with the robustness operator $OpRob(B_{i,par})$ and

10 a determining of the vector of parameters on the basis of the weighted spectral coefficients.

8. A method according to claim 7, wherein the robustness operator $OpRob(B_{i,par})$ verifies the following relationship:

15
$$OpRob(B_{i,par}) = \left\{ \max \left(0, 25 + \frac{B_{i,par} - P(B_{i,newmod})}{B_{i,par} + 2P(B_{i,newmod})}, 0 \right) \right\}^2$$

$B_{i,par}$ being the spectral coefficient and $P(B_{i,newmod})$ being a parameter dependent on the noise level of the new noise model having activated the transition in the frequency channel i .

9. A method of voice recognition according to claim 1, wherein the
20 reference space updating phase comprises the following operations, on the basis of the basic spectral coefficients each associated with a frequency channel i , each expressing the spectral energy of a basic frame obtained during the learning phase:

the determining of a robustness operator $OpRob(B_{i,base})$ for each basic
25 spectral coefficient $B_{i,base}$, this robustness operator expressing the confidence to be attached to the spectral coefficient $B_{i,base}$ with respect to the noise level of the new noise model in the same frequency channel i ,
the weighting of the basic spectral coefficients $B_{i,base}$ with the respective robustness operators $OpRob(B_{i,base})$, and

30 the preparation of the updated references with the weighted spectral coefficients.

10. A method according to claim 9, wherein the robustness operator $\text{OpRob}(B_{i,\text{base}})$ for the updating of the reference space verifies the following relationship:

$$\text{OpRob}(B_{i,\text{base}}) = \left\{ \max \left(0.25 + \frac{B_{i,\text{base}} - P(B_{i,\text{newmod}})}{B_{i,\text{base}} + 2P(B_{i,\text{newmod}})}, 0 \right) \right\}^2$$

$B_{i,\text{base}}$ being the basic spectral coefficient and $P(B_{i,\text{newmod}})$ being a parameter depending on the noise level of the new noise model having activated the transition, in the frequency channel i .

11. A method according to claim 9, in which the references are prepared on the basis of compressed basic spectral coefficients, wherein the method uses a conversion table to convert the compressed basic spectral coefficients into compressed and weighted basic spectral coefficients.

12. A method according to claim 11, wherein the conversion table contains the non-compressed basic spectral coefficients $B_{i,\text{base}}$ obtained by application of the reverse of the compression function to the compressed basic coefficients and wherein the method comprises:
 a determining of the robustness operator $\text{OpRob}(B_{i,\text{base}})$ for each of the non-compressed basic spectral coefficients $B_{i,\text{base}}$,
 a weighting of the non-compressed basic spectral coefficients $B_{i,\text{base}}$,
 a compression of the non-compressed and weighted basic spectral coefficients so as to obtain the compressed and weighted basic spectral coefficients.

13. A method according to claim 1, using, as references, a sequence of temporal frames corresponding to one or more words, this sequence of frames being identified by a series of vectors of parameters, these parameters being obtained by compression of spectral coefficients.

14. A method according to claim 1, using, as references, a sequence of temporal frames corresponding to one or more phonemes, this sequence of frames being identified by the center and the standard deviation of a Gaussian function, this center and this standard deviation depending on the parameters of the vectors of parameters of the frames.

15. A method according to claim 1, comprising a phase of noise-

105050" 6675860

suppression in the speech-containing temporal frames before the parametrization phase.

16. A system of voice recognition in a noise-ridden acoustic signal for the implementation of the method according to one of the claims 1 to 15, wherein the system comprises:

means to acquire the acoustic signal, digitize it and subdivide it into temporal frames,

a parametrization chain to translate the temporal frames containing speech into vectors of parameters in the frequency domain,

shape-recognition means with a reference space acquired during a learning stage, to compare the vectors of parameters coming from the parametrization chain with the references, so as to obtain recognition by the determination of a reference that most closely approaches the vectors of parameters,

means for modelling the noise to reiteratively prepare noise models, a new noise model replacing a current noise model,

means for detecting a noise transition between the new noise model and the current noise model,

means to match the parametrization chain with the noise of the new noise model having activated the noise transition,

means to update the references of the reference space as a function of the noise level of the new noise model having activated the noise transition.

17. A system of voice recognition according to claim 16, wherein the means used to update the references of the reference space comprise a first memory space to store the updated references, these updated references having to replace current references used for shape recognition before the detection of the noise transition, these current references being stored in a second memory space.

18. A voice-recognition system according to claim 16, comprising a memory space to store compressed basic spectral coefficients obtained from basic spectral coefficients each associated with a frequency channel i , these basic spectral coefficients each expressing the spectral energy of a basic frame coming from the learning stage, a conversion table to convert the compressed

09851433-050901

A'

basic spectral coefficients into compressed and weighted basic spectral coefficients, each weighted by a robustness operator $OpRob(B_{i,base})$ as a function of the noise level of the new noise model having activated the noise transition and of the basic spectral coefficient ($B_{i,base}$) to be weighted, these

5 compressed and weighted basic spectral coefficients being used for the updating of the references of the reference space.

19. A voice-recognition system according to claim 16, comprising means for noise-suppression in the speech-containing temporal frames before their parametrization.

095113.050904
T06050.0015260